

FIG. 1. Structural organization of the FGF-R2 gene and demonstration of IIIb and IIIc mutually exclusive splicing. (A) Organization of the FGF-R2 protein domains (top) and genomic gene arrangement of the region in which alternative splicing yields transcripts containing either the IIIb or IIIc exon and encoding the second half of the third immunoglobulin (Ig)-like domain. TM, transmembrane domain, TK, tyrosine kinase domains. The solid box represents a highly acidic domain, and the thick line indicates the IIIb- or IIIc-encoded portion of the protein. Shaded boxes represent exons, and solid lines represent introns, with intron sizes indicated. U and D indicate the exons upstream and downstream of these alternative exons, respectively. (B) Scale representation of the exons (solid boxes) and introns (solid lines) with regions of high (at least 90%) rat-human intron sequence similarity (shaded boxes). Also shown are regions FS and FL and their sizes. nt, nucleotide.

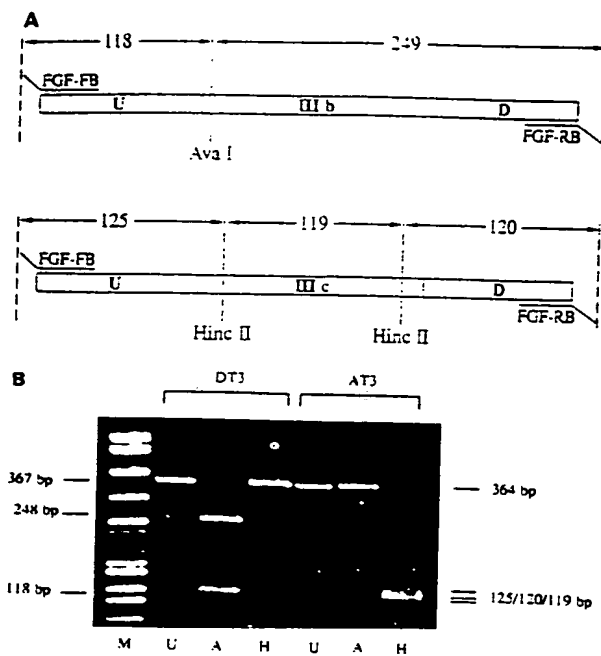


FIG. 2. Splicing of the endogenous gene transcript in DT3 and AT3 cells. (A) Map illustrating PCR products containing exon IIIb or IIIc amplified with primers FGF-FB and FGF-RB and sizes (in nucleotides) of fragments which result from *Ava*I or *Hinc*II digestion. U, upstream exon; D, downstream exon. (B) Gel showing the RT-PCR products following digestion with *Ava*I and *Hinc*II. DT3 cells express only products containing IIIb, and AT3 cells express products containing IIIc. U, uncut products; A, *Ava*I-digested products; H, *Hinc*II-digested products; M, pBR322 Msp I DNA size markers.

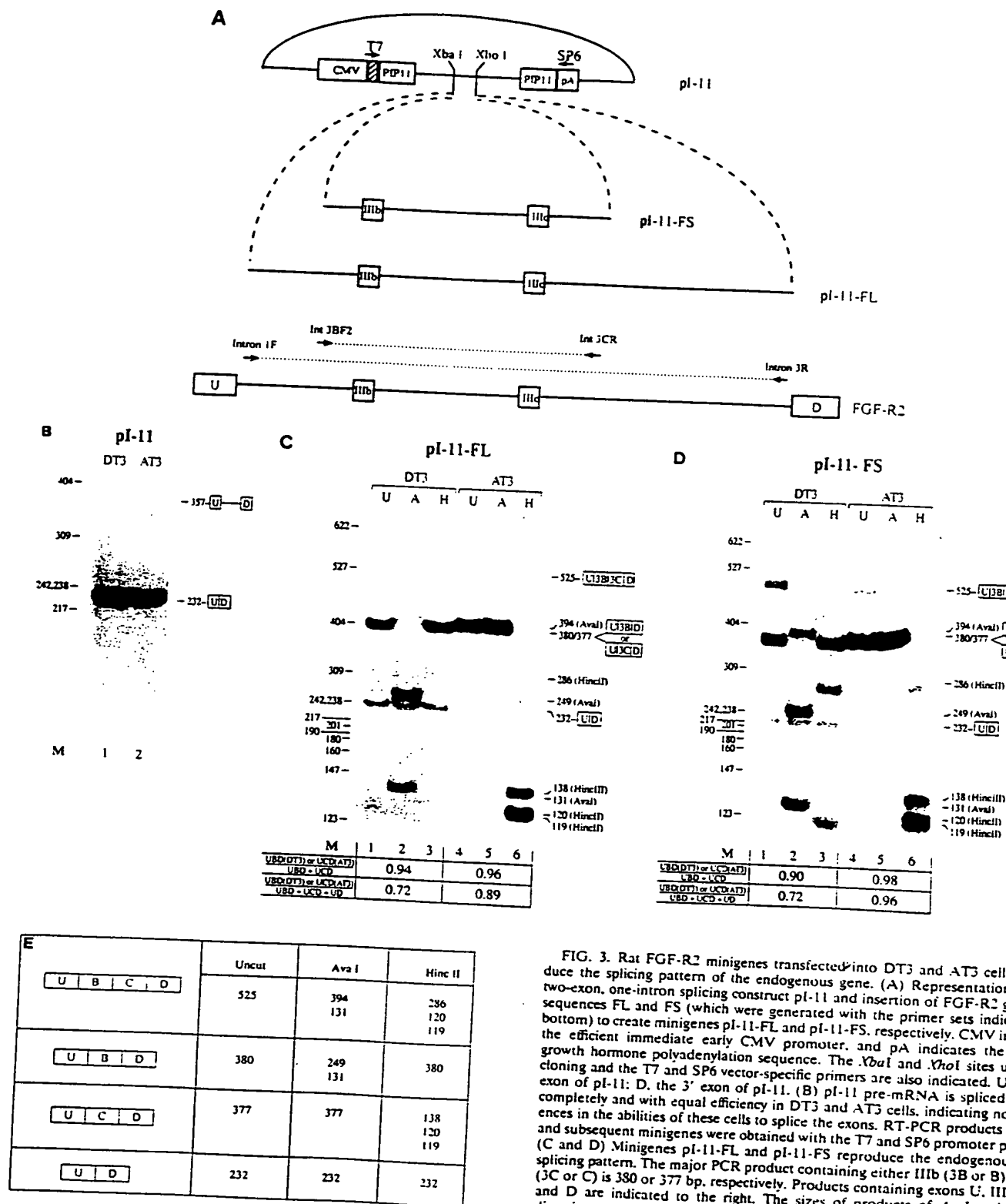


FIG. 3. Rat FGF-R2 minigenes transfected into DT3 and AT3 cells reproduce the splicing pattern of the endogenous gene. (A) Representation of the two-exon, one-intron splicing construct pl-11 and insertion of FGF-R2 genomic sequences FL and FS (which were generated with the primer sets indicated at bottom) to create minigenes pl-11-FL and pl-11-FS, respectively. CMV indicates the efficient immediate early CMV promoter, and pA indicates the bovine growth hormone polyadenylation sequence. The *Xba*I and *Xho*I sites used for cloning and the T7 and SP6 vector-specific primers are also indicated. U, the 5' exon of pl-11; D, the 3' exon of pl-11. (B) pl-11 pre-mRNA is spliced almost completely and with equal efficiency in DT3 and AT3 cells, indicating no differences in the abilities of these cells to splice the exons. RT-PCR products for this and subsequent minigenes were obtained with the T7 and SP6 promoter primers. (C and D) Minigenes pl-11-FL and pl-11-FS reproduce the endogenous gene splicing pattern. The major PCR product containing either IIIb (3B or B) or IIIc (3C or C) is 380 or 377 bp, respectively. Products containing exons U, IIIb, IIIc, and D are indicated to the right. The sizes of products of *Ava*I and *Hinc*II digestion are also indicated. Quantification was performed to yield values for the fraction of the expected IIIb (in DT3) or IIIc (in AT3) exon as a fraction of products containing IIIb and IIIc and also as a fraction of products skipping IIIb and IIIc (see Results and Materials and Methods). (E) Representation of the origins (in nucleotides) of the products obtained when UBD, UCD, and UBCD products are cut with *Ava*I and *Hinc*II. Sizes are indicated in base pairs. Lanes are labeled as in Fig. 2.

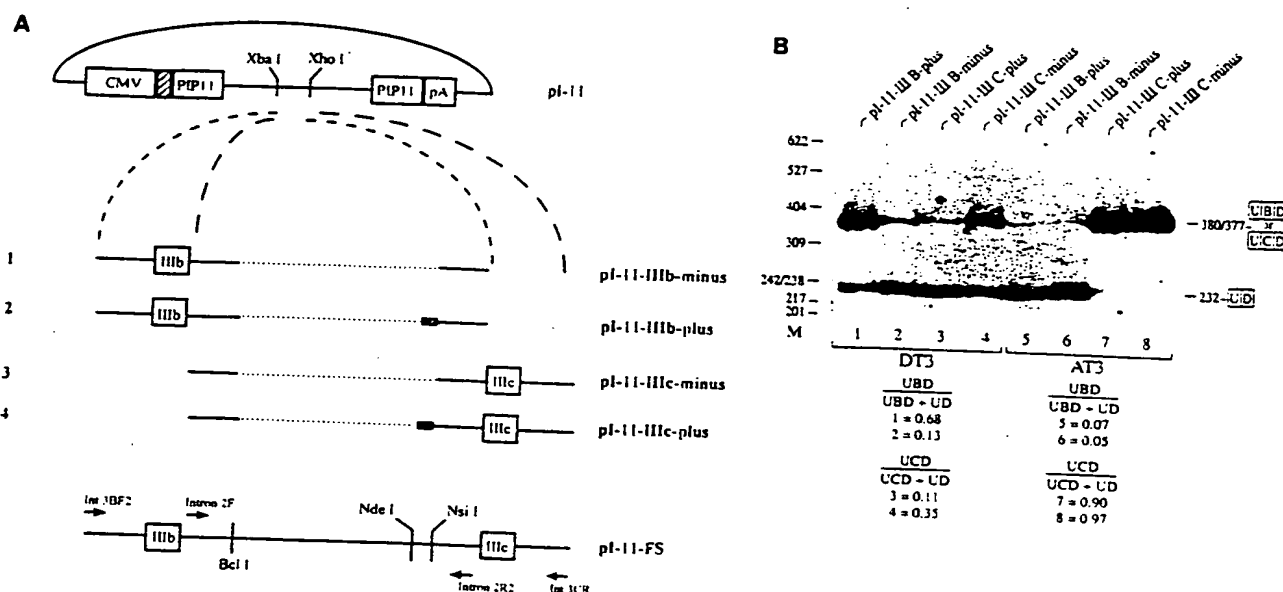


FIG. 5. Sequences contained between the *NdeI* and *NsiI* sites of intron 2 normally function to activate upstream IIIb splicing and repress downstream IIIc splicing. (A) Method used to generate minigene constructs containing either the IIIb or IIIc exon with *NdeI*-to-*NsiI* sequences (crosshatched boxes) present or deleted. All constructs had sequences *BclI* to *NdeI*, which were previously shown to be dispensable for regulation, deleted. The primers used to generate these regions in relation to the sequences of pi-11-FS are shown. The hatched box represents polylinker sequences present in PCDNA 3. (B) Transfection of the minigenes into DT3 and AT3 cells reveals that AT3 cells use exon IIIc highly efficiently and do not use exon IIIb efficiently regardless of the presence of *NdeI*-to-*NsiI* sequences. DT3 cells use exon IIIb efficiently only when these *NdeI*-to-*NsiI* sequences are present downstream. DT3 cells do not use exon IIIc efficiently, but when these sequences are deleted, IIIc usage triples. Quantifications were performed as described in Materials and Methods. Abbreviations are defined in the legend to Fig. 3.

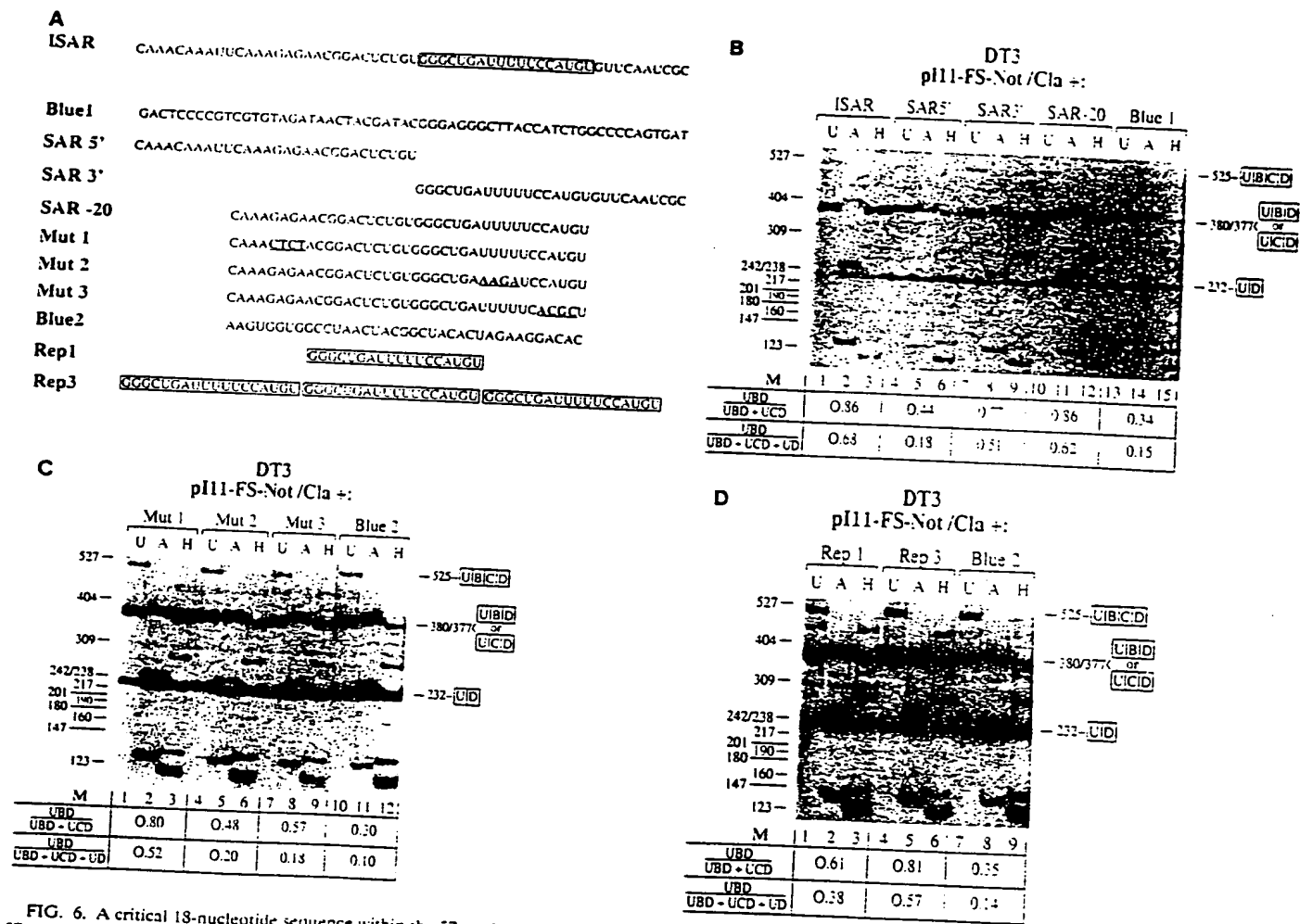


FIG. 6. A critical 18-nucleotide sequence within the 57-nucleotide ISAR sequence between *Nde*I and *Nsi*I nearly restores splicing regulation in DT3 cells. (A) The 57-nucleotide ISAR sequence is indicated at the top, and deletions and mutants of this sequence are shown below, as are control pBluescript sequences. The 18-nucleotide core sequence (Rep1) is boxed, and mutant sequences are underlined and in boldface. All sequences were tested by deleting ISAR sequences from p11-FS-Not/Cla-ISAR and inserting the indicated sequences. (B) SAR-20 and SAR 3' sequences restore regulation, whereas SAR 5' does not. (C) Mutations in the 18-nucleotide sequence shared by SAR-20 and SAR 3' (Mut2 and Mut3) cause loss of regulation, whereas a mutation outside this region (Mut1) preserves regulation. (D) One or three copies of the 18-nucleotide core sequence restore splicing regulation, with three repeats of the sequence being slightly more efficient than one repeat. Abbreviations are defined in the legends to Fig. 2 and 3.

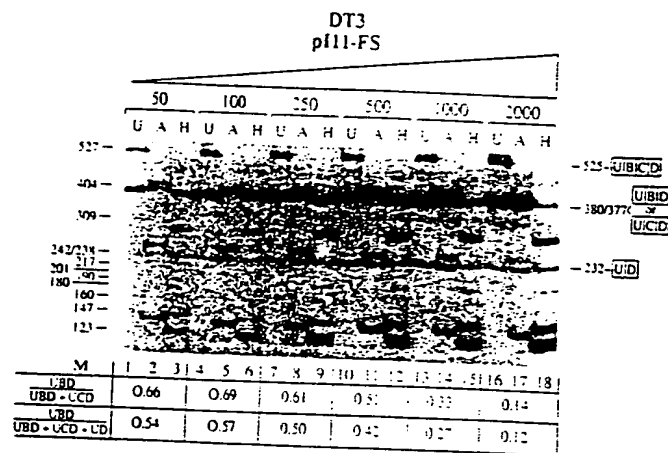


FIG. 7. DT3 cells contain a titratable factor or factors required for appropriate splicing regulation which can be overcome in transient transfections. Transient transfection of DT3 cells with increasing numbers of pI11-FS minigenes resulted in stepwise loss of IIIb inclusion and increased IIIc inclusion, suggesting that a factor or factors required for regulation (i.e., IIIb inclusion and/or IIIc exclusion) is overwhelmed when large numbers of these minigenes are transfected. Abbreviations are defined in the legends to Fig. 2 and 3.

A.

Rat	CCAUGGAAAAAUGCCACAAU
Human	CCAUGGAAAAAUGCCACAAC

B.

Rat	CAAA-CAAA-----UUCAAAGAGAACGGAC-UCUGUGGGCUGAUUUUU-CCAUGUUUUCAAUCGC
Human	CAAACCAAGCACAGGCCAAGAGAACGGACCUCUGUGGGUUGAUUUUUCCAUGCUUUUGAUUGC

FIG. 8. Intron sequences important for regulation of rat and human FGF-R2 splicing are highly similar. (A) Rat intron sequences corresponding to a previously reported 21-nucleotide human sequence, IAS2 (see Results), which also mediates IIIb activation, contain only 1 nucleotide difference. (B) The 57-nucleotide rat ISAR sequence is highly similar to human sequences in this same region, including the 18 nucleotides shown to be most important for regulation (boxed sequences).

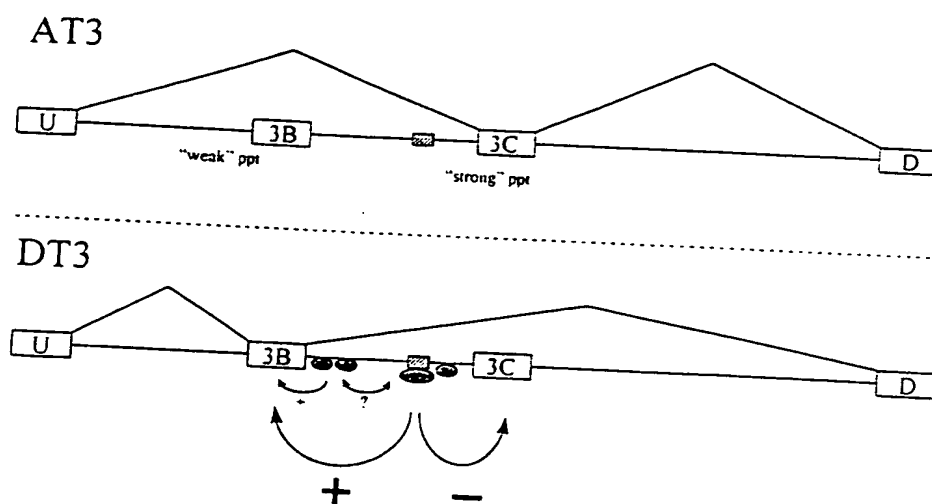


FIG. 9. Depiction of a model which can account for our results and the high fidelity of FGF-R2 splicing. AT3 cells use a default splicing pathway and choose the IIIc exon because of its stronger polypyrimidine tract (ppt); they splice IIIb inefficiently due to its weaker polypyrimidine tract. DT3 cells require a regulatory factor(s) which can activate (+) the weaker IIIb exon and at the same time repress (-) use of the IIIc exon. The ISAR element (indicated by a hatched box) is shown binding a factor or complex of factors (large shaded oval) which mediates both of these effects. The previously demonstrated contributions of other cis elements and associated factors (smaller shaded ovals) to IIIb activation are also shown, as well as the suggestion of possible cooperative interaction between proteins bound at several locations within the intron. Abbreviations are defined in the legend to Fig. 3.

GTAAC AACGTTTTTG TGTGTGTT
TTTTATTTT TATTTTATT TTTTTTTTA AGAAACTGA ATATAGGAGT
TAAAAAAGAC TCGGTGCTTT GGGAGGCAGC AGGCAGCTTC TAGAATAACT
CTTGTGGTCT TGGTATATTT ATAATGATCT TTCTTTGGTG GTGCAGCTGG
CGTCATGCCA GTGGCCATGG AAAAATGCCC ACAATGTTCA AAGTGCTTGA
AGATTATCTT CCACCCCCAC CCTGTTTTCA AGCCCTTCTT TCTGGTCTGT
CTTGTTTGGA CTGCACACTT CCCGTGATCA CTGTGTCTGA GTGCACGTGG
GCCTTGCGTT TGCATGCCCCG TCGAGTCTGC ACTCTCTGAT TATTAAGCCA
GACTTGGTTG CCTTTTATGC TAGTGACATA GAGAAATGCT AGCATGATAG
GATTCACCTA ACGAAAGTTT TGTTCTTTGG TTCGATTCCA CACCGGATCC
TTTCCAAAAC TGGAGAATGG TTATCTTCTA GTGCGTATGA CACTGGAGGA
TAGTGAAGGC AGATGGTGGG GTTTTCAGTT ATCATTCTTC ACACGCAGAC
ATATTCATAT TAGAAAAGGA AACAAACCAT AAATCCAGTT TTTTCTGTTA
CCAGTATTAC ACTTTCTGCC ATGTTCTTTC AATGATCATA TAAAGCAAGA
TGATTTTCGG CCTGAATGAA ATTAACCAGA ATCCAGTCAC CAAGATAAAG
TCCCACCCTG GTTCCCATGG AGCCTGAGGG ATGTGTGGGA TGTCCACCTG
ATCTGCCGTG CTTTATTCCA TCACACAGAA AATAGAAGAG CCTCCCCTTT
TCTCACAATT GGAGTCTGCA TCCAACAGGA CCAGAACCCA GATTAGCCCT
CAGGGTATTA TACTTTTGG AAACCCACTC CCAAATCCAT ATGCAAACAA
ATTCAAAGAG AACGGACTCT GTGGGCTGAT TTTTCCATGT GTTCAATCGC
ATGCATGTCT AAGGTGGTGA CGCCGGTGTG GTGATGGGCC TGCAGAGGTG
AGCTGGCCGG TGTCTCTCAG TGTCTCTTGG TTGTGGGCTT TGTGGACGGG
CTGCAGTTGG AATCTCCTGA TGGCCAGCAC CCCCTGGACC TGCTGGGACA
AGGCCTCTTG GTTCCAAGGC CCCCTCCACA ATCATTCTTA TGTCTAGCCT
TTTTCTTGCT TCGTTTGTTT TCTAG

Fig. 10

1 → GTAACAAT GCTTCATTTT TGTCTTTTTT TAAAAAGAAA GCTGGATATA

GAAGCTGAAA AGACTTGGTG CTTTGGGAGA CTGCAGGCAG CTTATAGGAT
 AACTCTTG TG GCCTTGGTAT ATTTATAATA ATCTTTCTTC GGTGATGCAG
 CTGGTATGAT GCCAGTAGCC ATGGAAAAAT GCCACAACG TTCAAAGTGC
 TTGCTCCAAT TTCTTCTAGA GATTAGCCTC CACCCCCACC CAGTTTTTAA
 GTTGTTCCCTT CTGGTTGATC TTGTTTAGGC TGCACATTTT CCATCATTAC
 TGCACATTAA CACCATTAA AACACACGCT TCCATGCCTG TTTAATACGG
 GGCATTTGAA TATCAGCAGA GTTTGTCCAA GTTTTAGGG AAATATTGGC
 AAGATGCAAT TTGTTCAACA AAGCATCATT TCTTTGGTTG CATGGTTGAT
 CCTTATGAGT TGCTGTTCTT GACCTTGTG CACCAAATTT GAGGGGAGCT
 CATCTTAATG AATGTACTAC TGGACGCTAC TAAAGGCAA AGGTTGACTT
 TTTAGGTTTG TCATGACTCA CATCCAAATG TTTATTAATG AAAAGAGAAA
 AAGCCCAGTT TTTTGGTTA CCAAGATGAT GCTTGCTTCC ATTTCTTTTT
 GTCAATGCTA TGTAGGGCAA GATGGTATCG CAGAAGTAAA AATAACCAGA
 GCCTGGTAAC CAAGACAACC TTCCACCCCA ATTGGTTCCC ACAGGGCCAG
 GAGGATGGGT GAGGTGCCCA TCTGGGCTTA TGTGCAGTGT GTTGTCTTAA
 AACACAGCAA TTTAGATAGA ACTACCCTTT CCTCTTGGTG GGAGTCTGCA
 GCCAACAGGA CCAGAACCAG CTTGGCCTTC TGGGCACCAT ACTTTTGGAA
 AACCACCCCT AAATGCAAAC CAAAGCACAG GCCAAGAGAA CGGACCTCTG
 TGGGTTGATT TTTTCCATGC GTTTGATTGC GTGCATGTGT AGGAGGTGAA
 GCCGGTGTGG TGACGGGCCT GTGGAGGTGA GCTGGTCAGT GTTGCTCCGT
 GTCTCTCGGT TGTGGGACTT TGTGGATGGG CTGCAGTCGG AATCTCCCAG
 TGGCCAGCAC CCCCTGAAGC CCCCAGTGCG ACGCCTTGTG GTTCCACAGC
 CCCCTCCACA ATCATTCCTG TGTCGTCTAG CCTTTCTTTT TGCTTCCCTT
 GTTTTCTAG

EGFR2 gene: ^{exons}
 Human intron between ¹III b + III c

Fig. 11